УДК 004.4
ББК 32.372
Х20

Харенслак Б., де Руйтер Дж.
Х20 Apache Airflow и конвейеры обработки данных / пер. с англ. Д. А. Беликова. – М.: ДМК Пресс, 2022. – 502 с.: ил.

ISBN 978-5-97060-970-5

Конвейеры обработки данных управляют потоком данных с момента их первоначального сбора до консолидации, очистки, анализа, визуализации и многого другого. Эта книга научит вас создавать и сопровождать эффективные конвейеры обработки данных с использованием платформы Apache Airflow.

Те, кто мало знаком с Airflow, получат базовое представление о принципах работы этой платформы в I части книги. Далее обсуждаются такие темы, как создание собственных компонентов, тестирование, передовые практики и развертывание, – эти главы можно читать в произвольном порядке в зависимости от конкретных потребностей читателя.

Издание предназначено для специалистов по DevOps, обработке и хранению данных, машинному обучению, а также системных администраторов с навыками программирования на Python.

УДК 004.4
ББК 32.372

ISBN 978-1-6172-9690-1 (англ.)
ISBN 978-5-97060-970-5 (рус.)

Оглавление

Содержание

ɧqdqɏɘɎ bɆ dəqɆbɎɜə d ɕɑɑbɔɏ ɟqdqɎqɏè

ɳqdqɣeɎ bɆ dɐdƎbɎɜə d ɕɑɑbɔɣ Ɉqdqɣqɣè

ɳqdqɏeɎ bɆ dedⱻbɎɜe d ɕɑɑbɔɏ Ɉqdqɏqɏè